

THE OHIO JOURNAL OF SCIENCE

VOL. XXVII

SEPTEMBER, 1927

No. 5

THE OBSERVER: AN INSTRUMENT OF PRECISION.

FRANCIS N. MAXFIELD

Department of Psychology, Ohio State University.

The development of methods and technique in measurement is fundamental to the progress of any science. A science, like physics, in which these methods have been developed to a high degree is referred to as an *exact science*. The tools of an exact science are referred to as *instruments of precision*. Mathematical treatment of the data of scientific observation gives the scientist various measures of the degree of precision obtained. It must always be kept in mind, however, that the idea of "exact" or of "precision" is relative rather than absolute.

In practical life as well as in making scientific observations, we are constantly making use of instruments of precision. I look at my watch to see what time it is. Compared with the dollar watch I carried recently while this watch was being cleaned, my own watch is an instrument of precision. Compared with my watch, the chronometer in the window of the shop by which I set my watch last week is an instrument of precision, and yet this chronometer is corrected from time to time on the basis of reports telegraphed from the U. S. Naval Observatory at Washington, where other instruments of precision,—transits, micrometers, verniers, and the like, are used to determine the correct time. Thus we see that precision is a relative rather than an absolute matter. To take an example from Merriam's "Method of Least Squares", a certain angle was measured by the same observer with three different instruments, a theodolite, a sextant, and a transit. Four observations with the theodolite gave a result of $6^{\circ}17'5''$, with a probable error of 1.4". Five observations with a sextant gave $6^{\circ}17'12''$, with a probable error of 5.8". Six observations with the

transit gave $6^{\circ}17'0''$ with a probable error of $23.3''$. The true value of the angle is certainly in the neighborhood of $6^{\circ}17'$, since the results of all three instruments agree in this. But the limits between which the true result may be expected to lie vary with the three instruments. The true value as determined by the theodolite, the most accurate of the three, may, with a probability greater than one-half, be considered to lie between the values of $6^{\circ}17'5''-1.4''$ and $6^{\circ}17'5''+1.4''$, that is, between $6^{\circ}17'3.6''$ and $6^{\circ}17'17.8''$; and the transit, the least accurate of all, gives limiting values of $6^{\circ}16'36.7''$ and $6^{\circ}17'23.3''$. These are all instruments of precision and yet one has a higher degree of precision than another and in no case is the precision absolute.

Now when we consider the methods used in astronomy, physics, or any other of these so-called exact sciences we find that instruments of precision play an important part in them all, both in the observation of natural phenomena and in the development of experimental technique, as well as in the standardization or correction of other instruments. We realize further that in any of these methods the observer, the man who looks through the telescope or reads the thermometer, plays an important part. In fact, for the purposes of science, the observer himself is an instrument of precision. It is, therefore, interesting to the worker in any field of science to make a study of this useful and necessary instrument. The study of this human instrument is a major problem in the science of psychology.

The development of modern psychology has been characterized by a breaking away from the psychology of the past, which was based deductively upon a speculative philosophy, and by the foundation of a new experimental science, bringing to its service many of the instruments and methods of the older sciences, and inventing instruments and methods of its own. This modern movement is usually dated from the year 1878, when Wundt (1832-1916), then holding the chair of philosophy at Leipzig, founded at that university the first laboratory of experimental psychology. Wundt had begun his career as a professor of physiology at Heidelberg and brought to the new science of psychology the data and experimental procedure of this older science. In the nearly fifty years since that time, psychology has been developed as an experimental science which can no more stand by itself without dependence upon physiology than modern physiology can exist without chemistry.

Though the object of psychology continued to be defined as the study of the processes of consciousness, the principle came to be universally admitted that nothing happens in consciousness which is not conditioned by certain physiological processes. If one accepts the definition of psychology given by the more recent behavioristic school that "psychology is the science of the reaction of the individual to the stimuli to which his environment subjects him", this notion of the observer as an instrument of precision is still further emphasized.

Accepting this definition, we find that we are interested in an individual human being, not as a mind or soul made up of or possessing certain faculties such as cognition, will, and reason, but as an organism that reacts in a certain manner in a certain situation. We find that this reaction is determined by the arousal of certain physiological processes involving particularly the nervous system of the organism, and we have consequently seen the development of a psychology of the lower animals and of the human infant as well as an objective study of the human adult. It matters little for our purposes whether we consider only the reactions of the observer in a scientific experiment, or also make reference to his conscious states, and the terms "perceive" and "perception" will be used here without any necessary implication of one point of view or the other. In the sense that an instrument of precision is a *mechanism*, it may appear that the argument is in favor of a behavioristic interpretation, but the discussion is just as, or perhaps even more, pertinent to the methods of scientific observation if one makes his interpretations from the standpoint of psychophysical parallelism or interactionism.

That branch of psychology which seeks to determine the functional relation between the physical processes which we call stimuli and the reactions or mental processes of the organism is called *psychophysics*. For example, I ask someone to tell me which of the two weights is the heavier. If one of these weighs 100 grams and the other 120 grams he will very readily select the heavier, whereas, if one of them weighs 100 grams and the other 101 grams, he is very likely to tell me that the two are equal. Now the first weights act as stimuli giving rise to certain sensory responses or perceptions of weight due to strain upon the muscles of his arm which are richly supplied with sensory nerve endings. Certain neural and brain processes are involved. The second case is similar, but in the first case the

weights are perceived as different and the judgment "heavier" is given, and in the other "equal", by which we imply that his perception of weight is not different from that aroused by the standard weight with which it is compared. We are interested in determining under what conditions his reactions to the weights are the same and under what conditions they are different. We seek further to formulate some mathematical statement which will generalize our findings, just as the astronomer or physicist derives his mathematical formulae.

We have here the concept of the sensitivity of the observer just as we speak of the sensitivity of a chemical balance which is used in comparing a given object with certain standard weights or masses. Closely connected with this idea is the concept of the *threshold*, a term introduced into psychology by Herbart in 1811. It is perfectly obvious that, in order to be perceived at all, a stimulus must have a certain intensity. Suppose we are looking at the stars on a clear moonless night. We can see stars up to the first, second, and third magnitude, and if our eyesight is keen, up to the sixth magnitude. But we know from using a telescope that there are thousands of stars too faint to be seen by the unaided eye. The intensity of light from one of these is not great enough to arouse a sensory response. In other words, its intensity is below the threshold. The threshold may be defined, therefore, as that value of the stimulus which is just sufficient to produce a sensory response, less values producing no response.

It is clear that the value of this threshold for any given kind of stimulus may vary for different individuals and in the case of vision, for example, for the two eyes, and for different parts of the retina of the same eye, in the same individual. Its value will also vary with varying conditions of attention, expectation, practice, and fatigue. We see at once that the value of this absolute threshold of intensity is an index of the sensitivity of the observer. The lower the threshold, the greater his sensitivity; and the higher this threshold, the less his sensitivity.

To take another illustration from astronomy. Two observers, A and B, are gazing at a small constellation. A is a trained astronomer, while B is an amateur. A sees nine stars in the constellation, while B sees only seven. A makes a drawing of the stars in this constellation and shows B where the two stars which the latter does not see should appear, or he allows

him to observe them through a good field glass. He further shows B how to focus his eyes on a point a little to one side of the position where the stars are to be seen, as for dim light the center of vision is not the most sensitive portion of the retina. These changes in the method of observation, this practice, and attentive expectation act to lower B's threshold and he finds that he can now see nine stars in the constellation just as A does. His absolute threshold has been lowered and his sensitivity increased.

We sometimes speak of an upper threshold or limit of sensibility in contrast with that just described which is designated as the lower threshold. The upper threshold is that stimulus value beyond which there is no sensory response. Both are well illustrated in the case of tones produced by vibrating bodies or air columns. Some observers have claimed to hear as a tone the vibrations of a tuning fork vibrating sixteen times per second, but for most of us this lower limit is about thirty-two vibrations per second. On the other hand, as will be observed if we listen to a Galton whistle, we may shorten a vibrating air column, making the pitch higher and higher, until we pass the upper limit of sensitivity and no tone is perceived. This upper threshold for pitch is about 36,000 vibrations per second.

In light waves we find a similar phenomenon. If a beam of sunlight is diffracted by a prism or grating and thrown on a screen, we get a spectrum, but only a part of this will be visible to the human eye. We see violet, blue, green, yellow, orange, red; but it can be easily demonstrated in the physics laboratory that beyond the violet are the ultra-violet rays and beyond the red are the infra-red rays. The former are beyond the upper and the latter below the lower limit of sensitivity or wave length threshold and the latter below the lower limit of sensitivity or wave length threshold for the human retina.

The lower threshold values in animals are often quite different from those observed in the human species. Dogs have a remarkable sensitivity to olfactory stimuli. Romanes tells the story of a dog whose olfactory sensitivity was tested in the following manner. The dog's master walked nearly across a large field and turned abruptly to the right. He was closely followed by twenty-four men, each of whom tried to step exactly in the footsteps of the man in front of him. At the place where the owner of the dog turned to the right, the first

man following turned to the left, the second to the right and so alternately until twelve had followed to the right and the other twelve had gone to the left, each stepping in the other's tracks as before. A short time afterward, the dog was set on his master's trail, which he followed rapidly to the turning point. Here he ran past, but returned to "pick up the scent". This he did readily, following his master's trail without hesitation.

The problems of the psychophysicist are not limited to the determination of the limits of sensitivity—these upper and lower thresholds. Lying between these limits are series of stimuli all of which may give rise to parallel series of sensory responses varying in quality, intensity, and duration. The stimuli consist of some form of energy—kinetic energy, light, heat, etc.—which may be measured in terms of physical units. This light is 32 candle power, that is 16; this is a weight of 100 grams, that of 120, etc. Roughly speaking, the stimulus intensity series and the response intensity series are parallel. Two weights are heavier than one. A 32-candle-power light is brighter than one of 16 candle power. So true is this parallelism that our language is often confusing. We can not be certain which series is meant when some one says that "one sound is louder than another." The statement may refer to response intensity or to the intensity of the sound waves of the stimulus, or to both.

In spite of this rough parallelism, it can easily be shown that the response intensities may be indistinguishable, whereas the stimuli are clearly different, or that stimuli of equal intensities may be perceived as different. It is, therefore, the problem of psychophysics to study the relation which does exist between these two series.

This brings us to the concept of the *difference threshold*. In lifting two weights we may not be able to discriminate between 101 grams and 100 grams, the judgment "equal" being given. It is obvious that if we gradually add to the 101 gram weight we shall finally get a weight which will be judged "heavier" than our 100 gram weight. If such a weight is 103 grams, our difference threshold in the direction of increase is 3 grams. That is, a difference of 3 grams in stimulus intensity is just sufficient to produce discrimination. In a similar way by comparing weights of 99, 98, 97 grams, etc., with our original standard weight of 100 we come to a weight which is just perceptibly lighter than our standard. If this is 98 grams, our

difference threshold in the direction of decrease is 2 grams. A more sensitive subject might distinguish 102 and 99 grams from 100. His difference thresholds in this case would be therefore 2 grams in the direction of increase and 1 gram in the direction of decrease. The interval between these two difference thresholds is known as the *interval of uncertainty* since the person lifting the weights included in this interval can not judge better than by chance whether the comparison weight is heavier or lighter than the standard.

These measures of difference have been called "just perceptible differences." By taking account of these just perceptible differences in any given stimulus scale, either of quality, intensity, duration, or extent, we may arrange stimulus series which are paralleled by series of just distinguishable sensory responses or perceptions. The difference threshold may, therefore, be defined as the smallest different in two stimuli such that the two will be discriminated better than by chance. As it is a matter of relative stimulus value it may be stated in terms of the units of stimulus value or as a percentage. In this sense the difference threshold is comparable to the concept of probable error, the interval of uncertainty corresponding to the interval between the plus and the minus probable error.

The series of just distinguishable sensory responses are often very extensive. For example, in the series in the octave between A^1 with 435 vibrations per second and A^{11} with 870, there are over 1200 tones which can be discriminated or are "just noticeably different", in the practiced ear of a musician.

It will make the concept of the lower threshold, or stimulus *limen* and that of the difference threshold, or difference *limen* clearer to mention a few analogies. The first is that of the tangent galvanometer first suggested by Delboeuf. (1831-1896).

The tangent galvanometer is an instrument for measuring an electric current by means of a magnetic needle suspended in the field of a coil of wire through which the current to be measured is passed. The angular deflection of the magnetic needle is observed and read from a graduated scale. It is found that the greater the current passed through the coil, the greater the total deflection of the needle from its original position. It is further observed that the amount of angular deflection due to a given increase in current is not, however, directly proportional to the increase in current, these increases producing less and less additional deflection the farther the

needle swings from the original position. Moreover, no current can be made strong enough to cause a deflection of 90° . It has been found that the current intensity is directly proportional to the tangent of the angle of deflection. The instrument is, therefore, called a tangent galvanometer.

If we observe the action of this instrument as our electric current is gradually increased we find an interesting phenomenon which is analogous to that observed in our stimulus-response series. First we note that the current has to be of a certain strength before the inertia of the needle and the torsive resistance of the suspension fiber are overcome and any deflection at all is noted. This corresponds to our lower limit of sensibility, our lower stimulus threshold. If a finer needle with a gossamer suspension fiber is used, this resistance may be reduced and our instrument rendered more sensitive, i. e. its stimulus threshold has been lowered.

The analogy can be carried still further, in that, when the needle is held at any given angle by a current of a certain strength, a further deflection will not take place, for the reasons mentioned above, until a certain amount of increase of current has been made which can overcome these resistances. This necessary increase is analogous to the difference threshold and may be greater or less according to the sensitivity of the instrument and to the position of the needle.

Another convenient analogy may be found in the chemical balance. Here scale pans are mounted at the two ends of a horizontal beam swinging freely in the vertical plane on an axis perpendicular to this plane of rotation. If these pans are in equilibrium a tiny bit of dust falling on one of them will not cause a tilting of the beam. But if dust should be allowed to accumulate on one of the pans the resistance due to inertia and friction would finally be overcome and this scale pan would drop and the apparent equilibrium would be destroyed. The mass of the dust necessary to effect this disturbance of equilibrium may be compared to the lower threshold of sensibility, the just perceptible stimulus already mentioned. If the balance were made more sensitive by reducing the friction, or the mass of the pans, or by increasing the length of the beam, this equilibrium would be disturbed by a smaller amount of dust, i. e., its stimulus threshold would be lowered.

If two equal masses are placed on these scale pans in their previous condition of equilibrium they will balance one another

and we have again a condition of equilibrium. To disturb this again it is necessary to add an appreciable amount to one of the pans. This is comparable to the difference threshold or just perceptible difference.

A third illustration may be found in observing the elasticity of a rubber band. It is necessary to exert a certain amount of force upon it before it will stretch at all. Again we have our lower threshold or limit of sensibility. After it has begun to stretch, relative changes in length will be proportional to the force exerted upon it. If held in a stretched position, however, an appreciable increase in the force exerted must be made before it will stretch further. Again we have our difference threshold. If force enough is applied, the limit of elasticity will be reached, the rubber band will break, and we have an analogy to our upper threshold or upper limit of sensitivity.

The human subject may be compared to such an instrument as the chemical balance. A weight must be of a certain intensity before it is perceived at all. A weight held in the other hand must have a "just noticeable difference" added to it before it is perceived as heavier. Here we have the stimulus threshold or limit of sensitivity and the threshold of difference.

These thresholds may be defined in specific units such as grams, candle-power, amperes, etc., or they may all be reduced to the absolute unit or energy, the erg. Thus Langley determined that the just perceptible light sensation under favorable conditions was stimulated by .0000003 erg. The just perceptible sound stimulus is represented by a figure considerably smaller.

Interest in these quantitative aspects of psychology, or in the science of psychophysics as it came to be called, dates from the researches of the German physiologist, Ernest Heinrich Weber (1795-1878), at the University of Leipzig, who in 1849 published his celebrated work "On the Sense of Touch and Organic Feelings." He experimented with lifted weights and found that he could just distinguish weights of 32 and 35 drachms. He found further that in order to make a weight just noticeably different from one of 32 ounces it was necessary to make a certain proportional rather than an absolute difference between the weights compared, i. e. 3 ounces instead of 3 drachms. Weber stated his conclusions as follows: "In the discrimination of objects that are compared the one with another, we do not perceive the difference between the objects but the ratio of this difference to the magnitude of the compared objects."

His work was followed by that of Gustav Theodor Fechner (1801-1887), who had at one time held the chair of physics and later that of philosophy at Leipzig. Fechner gave the name "psychophysics" to the new field of investigation, carried on many experiments in practically all the different sense fields, and wrote extensively on the subject's mathematical and philosophical aspects. He gave the name of Weber's Law to the psychophysical principle as stated by his predecessor.

Fechner sought to give a mathematical statement to the psychophysical law. The functional relationship between the sensory response series and the accompanying stimulus series he expressed by grading the former in an arithmetical progression and the latter in a geometrical progression. The mathematical relation between two such progressions or series may be expressed by saying that successive terms of one are respectively directly proportional to the logarithms of the successive terms of the other. Fechner's mathematical statement of the psychophysical law then becomes—

$$R \propto \log S$$

where R is the response or sensation and S the stimulus. Here we have a formula analogous to that for the magnetic deflection of the tangent galvanometer where

$$A \propto \tan \theta$$

when A equals the current strength and θ equals the total angle of deflection.

Fechner's law, as the logarithmic statement of the psychophysical principle has been called, may then be stated in one of two ways:

(1) *If stimuli are arranged in a geometrical series the sensations (responses) accompanying them will form an arithmetical series; or*

(2) *The sensation (response) is directly proportional to the logarithm of the stimulus.*

For Fechner, of course, the sensation aroused by a just noticeable stimulus was a sensation unit, in terms of which other supraliminal sensations and sensation differences as elements in consciousness could be measured.

An approximately exact illustration of this law is found in the magnitude of the visible stars. The classification of stars according to their brightness goes back to Hipparchus (125 B. C.) and Ptolemy, who divided the naked-eye stars arbitra-

rily into six "magnitudes", the first magnitude stars being some twenty of the very brightest and the sixth magnitude stars those just visible to the unaided eye. Now a first magnitude star is about one hundred times as bright as a sixth magnitude star. Stars may, therefore, be divided arbitrarily on the basis of photometric findings by the use of a light ratio of 2.51 ($\sqrt{100}=2.51$). That is, a star of the second magnitude is 2.51 times as bright as one of the third, one of the third 2.51 times as bright as one of the fourth, and so on. This method, first proposed by Pogson in 1850, has been generally adopted and most of the naked-eye stars have been measured photometrically and placed on this "absolute scale."

It would be gratifying if all stimulus series and response series parallel to them were found to conform to this simple functional relationship. When we test these series in different sensory fields for varying intensities, qualities, durations, etc., we find only a rough approximation of Fechner's law and that is usually limited to rather narrow ranges of these series. Weber's statement of the psychophysical principle is more nearly accurate because more general. We do not perceive or react to the difference between compared stimuli or objects, but the ratio of their difference to their magnitudes. Thus formulated we have a law which seems to be substantiated by the results of experimentation.

The data of extensive experimentation in this field both by the so called *gradation methods* and *error methods* (method of mean error, method of constant stimuli, etc.) are conveniently treated by the mathematics of probability and error in ways similar to the treatment of a series of data obtained from any scientific instrument of precision. The results show the sensitivity or degree of precision in any single observer as well as differences among different observers.

The statement of the principle of relativity as found in Weber's psychophysical law has not been confined to psychology. We find it in the principle of marginal utility in economics. As pointed out by Bernoulli (Daniel Bernoulli, 1700-1782), "we may regard the satisfaction which a person derives from his income as commencing when he has enough to support life, and afterwards as increasing by equal amounts with every equal percentage that is added to his income; and vice versa for loss of income." (A. Marshall, "Principles of Economics," 6th ed. p. 135.). Of course, the term "satisfac-

tion" is a psychological concept and the principle of Bernoulli and that of Weber are probably fundamentally the same. In other words, the "marginal utility" of any commodity to its possessor diminishes with every increase in the amount of this commodity which he already has.

Just as we notice the increase in light when a second electric light is turned on in a room but do not notice an additional light in the room already illuminated by a thousand lights, so ten dollars makes an appreciable increase in the monthly wages of a chauffeur but would not be noticed if added to the income of the millionaire whose car he drives. It is a "drop in the bucket" to use a folk-phrase expressing this principle, while "an inch on the end of a man's nose" is another folk-phrase to express a difference which is readily appreciable.

An example of this principle comparable to the psychophysical law is seen in the relation of rainfall to agricultural production. Before wheat can be grown at all a minimum amount of annual rainfall is necessary. This is the rainfall threshold, and we have a "just possible" wheat production. Even small increases in this rainfall yield rapid increases in the amount of wheat per acre which can be grown. But this rate of increase in production will fall off rapidly, and whereas an increase from 12 to 20 inches may double the crop, the increase from 40 to 48 will not do so. Finally a point is reached where further increase in rainfall will not increase the crop but may even be harmful and lessen production.

Scientific instruments frequently yield data which may show constant as well as variable errors. The constant error of a ship's chronometer is a certain regular rate of gain or loss per day. The instrument is set for Greenwich time and its time readings are compared from day to day with the true time as determined by some observatory. Its regular gain may be three seconds per day. This is a constant error. But continued observations show that it is not exactly three seconds: it is sometimes slightly more or slightly less than this. This difference is the variable error and may be due to one or many unavoidable causes. The navigator makes allowance for the constant error, but the variable error remains and must be taken into account as affecting the certainty of a given time observation. Considered as an instrument of precision the human being acting as an observer may show both constant and variable errors. This constant error when present in

observers is known as the "personal equation." The history of its discovery is interesting.

In the records of the Astronomical Observatory at Greenwich we find the following entry made by the British Astronomer Royal, Nevil Maskelyne, (1732-1811), who writes after holding that position for thirty years:

"I think it necessary to mention that my assistant, Mr. David Kinnebrook, who had observed the transits of stars and planets very well in agreement with me all the year 1794, and for the greater part of the present year, began from the beginning of August last to set them down half a second of time later than he should do according to my observations; and, in January of the succeeding year he increased his error to eight-tenths of a second. As he had unfortunately continued a considerable time in this error before I noticed it, and did not seem likely to get over it and return to a right method of observing, therefore, though with reluctance as he was a diligent and useful assistant to me in other respects, I parted with him.

"The error was discovered from the daily rate of the clock deduced from a star observed on one of two days by him and on the other by myself, coming out different to what it did from another star observed both days by the same person, either him or myself. . . .

"I cannot persuade myself," he continues, "that my late assistant continued in the use of this excellent method (i. e., Bradley's eye and ear method) of observing, but rather suppose he fell into some irregular and confused method of his own, as I do not see how he could have otherwise committed such gross errors."

This record made by Maskelyne is the first account we have of any observation of a personal equation. Maskelyne considered that the discrepancies between his and his assistant's observations were due to some faulty method on his assistant's part and did not concern himself with investigating the matter further. The incident of Mr. Kinnebrook's dismissal was, however, mentioned in a history of the Greenwich observatory published twenty years later (1816) and here it attracted the attention of Bessel (Friedrich Wilhelm Bessel, 1784-1846), the Königsberg astronomer.

It was hard for Bessel to see how an assistant, who would have every reason for bringing his observations into agreement with those of his superior, should have so persistently shown this constant error. To test the matter for himself he compared his own results with those of other astronomers. In December, 1820, he observed ten stars on alternate nights with Dr. Walbeck, determining the rate of the clock as Maskelyne and his assistant had done. When they first compared their results

they found a difference not of 0.8 seconds, but of 1.1 seconds. Being trained astronomers, they took particular precautions on the following two days, but with very little difference in the result, for the discrepancy was nearly one second (average for eight days 1.04 seconds). So careful were they that Bessel wrote, "We ended the observations with the conviction that it would be impossible for either to observe differently, even by only a single tenth of a second." Still Bessel was not satisfied. Walbeck was less experienced in transit observations than he, and so similar comparisons were made either directly or indirectly with many of the best astronomers of Europe.

It became recognized, therefore, that there is in each observer a tendency to observe star transits in a characteristic way which may differ in time results from other observers equally well trained (the relative personal equation), and that these time results differ by a more or less constant amount from the true time measured (the absolute personal equation). What applies in observing star transits by the eye and ear method applies more or less generally in all cases where a person reacts in any way to a certain stimulus. This "constant error" is analogous to the error of the chronometer and in a similar way it may be determined within certain limits of variation and allowed for. We see, therefore, that in many ways the observer is an instrument of precision, having, as do other scientific instruments of precision, a certain degree of sensitivity, a certain interval of uncertainty, and certain constant and variable errors. We may infer also that the analogy of the galvanometer, the chemical balance, etc. is not certainly merely an analogy.